

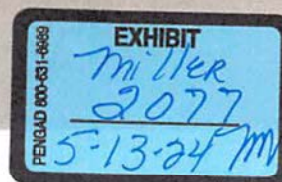
PRIMER ON

Multiple Sclerosis

BARBARA S. GIESSER

Merck 2077
Hopewell v Merck
IPR2023-00480

OXFORD



PRIMER ON MULTIPLE SCLEROSIS

EDITED BY

BARBARA S. GIESSER, MD, FAAN

Department of Neurology
David Geffen School of Medicine at UCLA
Los Angeles, CA

OXFORD
UNIVERSITY PRESS

2011



Oxford University Press, Inc., publishes works that further
Oxford University's objective of excellence
in research, scholarship, and education.

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi

Kuala Lumpur Madrid Melbourne Mexico City Nairobi

New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece

Guatemala Hungary Italy Japan Poland Portugal Singapore

South Korea Switzerland Thailand Turkey Ukraine Vietnam

Copyright © 2011 by Oxford University Press, Inc.

Published by Oxford University Press, Inc.

198 Madison Avenue, New York, New York 10016

www.oup.com

Oxford is a registered trademark of Oxford University Press

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
electronic, mechanical, photocopying, recording, or otherwise,
without the prior permission of Oxford University Press.

Library of Congress Cataloging-in-Publication Data

Primer on multiple sclerosis / edited by Barbara S. Giesser.

p. ; cm.

Includes bibliographical references and index.

ISBN 978-0-19-536928-1 (alk. paper)

1. Multiple sclerosis. I. Giesser, Barbara S.

[DNLM: 1. Multiple Sclerosis. WL 360 P953 2011]

RC377.P747 2011

616.8'34—dc22

2010000644

ISBN-13 9780195369281

The science of medicine is a rapidly changing field. As new research and clinical
experience broaden our knowledge, changes in treatment and drug therapy occur.
The author and publisher of this work have checked with sources believed to be
reliable in their efforts to provide information that is accurate and complete, and in
accordance with the standards accepted at the time of publication. However, in
light of the possibility of human error or changes in the practice of medicine, nei-
ther the author, nor the publisher, nor any other party who has been involved in the
preparation or publication of this work warrants that the information contained
herein is in every respect accurate or complete. Readers are encouraged to confirm
the information contained herein with other reliable sources, and are strongly
advised to check the product information sheet provided by the pharmaceutical
company for each drug they plan to administer.

9 8 7 6 5 4 3 2 1

Printed in the United States of America

on acid-free paper

29 Issues in the Design and Interpretation of Multiple Sclerosis Clinical Trials

Stephen Krieger, Svenja Oynhausen, and Aaron Miller

Multiple sclerosis (MS) is a disease whose heterogeneity poses unique challenges in making the diagnosis, offering prognosis, and deciding about treatment. The heterogeneity may pose even greater challenges in the design of clinical trials because it leads to problems of operational definitions, ascertainment of clinical data, and selection of meaningful outcomes as they pertain to characterizing the disease course. Applying the results of clinical trials to individual patients adds an additional degree of difficulty.

The natural history of MS has been well characterized over the past several decades. Although there are numerous methodological problems with the direct use of natural history controls, the entire enterprise of designing clinical trials for MS begins with applied natural history. Assumptions about the expected behavior of the disease are implicated in trial design, outcome selection, entrance criteria, and power calculations. Clinical trials of MS treatments are typically short term, relapse, or magnetic resonance imaging (MRI)-based studies; long-term benefits assessed utilizing robust clinical measures remain to be definitively established. As the disease course typically spans several decades, it is particularly difficult to draw firm conclusions about the consequence of treatments that have been available for only a fraction of this duration. Indeed, it is not clear how best to determine whether, and to what degree, current medications are influencing the long-term course of the disease (Noseworthy, 2007).

A discussion of MS clinical trial design and interpretation must begin with a critical review of the operational definitions used to characterize the disease. The broad range of MS disease course has

been delineated into four subtypes, relapsing-remitting MS (RRMS), secondary progressive MS (SPMS), primary progressive MS (PPMS), and progressive-relapsing MS (PRMS). It should be emphasized that the names for the clinical subtypes are of limited, mostly descriptive applicability. As they are based solely on the effects of the disease that cross the clinical threshold, the categories do not necessarily reflect the true underlying pathological heterogeneity. In RRMS, for instance, the formation of new T2 lesions is far more common than the occurrence of clinical attacks, indicating that even during periods of clinical quiescence, tissue damage continues to accumulate. The subtypes also vary in their definition as they apply to the temporal course of MS: PPMS is a discrete subtype, but RRMS and SPMS can both occur in the same individual at different points in his or her disease course. In addition, the transition from RRMS to SPMS is indistinct and can only be definitively identified in retrospect. One, therefore, cannot know whether a patient with RRMS has already begun to progress at the time of enrollment into an RRMS trial.

Nonetheless, these categories are most useful in the context of clinical trials, where homogenous populations are desirable to most clearly discern a therapeutic effect, and much of the successful work in the field, as well as the focus of this chapter, pertains to relapsing-remitting disease. Although necessary from a trial design perspective, the use of categories that are not biologically defined imposes several assumptions on the planning of a trial. As Randy Schapiro (personal communication) has noted, "There is no relapsing-remitting MS or secondary progressive MS—there is only MS."

While MS experts debate whether MS is one disease or many, confining clinical trials to a particular disease state not defined by distinct pathophysiologic mechanisms may increase both false-positive and false-negative results for clinical research. The reliance on these classifications for clinical trials limits the generalizability of the results across the entire MS spectrum, and it typically restricts approval and licensure of an agent to the subtype of MS in which it has been studied. An era in which the subtype-delineations are likely to be updated, as genetics and biomarkers become available to better elucidate the pathological substrates for clinical patterns, is likely beginning.

Clinical trials in MS have, since the early 1980s, followed a traditional “double-blind, placebo-controlled, randomized paradigm” (McFarland and Reingold, 2005) and have led to the approval of six agents for the treatment of MS. The widespread use of these treatments has transformed the management of MS and has significantly impacted the design of clinical trials that are needed to find safer and more effective therapies for relapsing MS and to test new therapies for other as yet untreatable forms of the disease (McFarland and Reingold, 2005). Despite the extraordinary advancements in neuroimmunology, rational drug development, and clinical trial design and analyses, clinical trials are hampered by an incomplete basic understanding of the MS disease process, the mechanism of action of the agents under investigation, and the ideal way to gauge their clinical effectiveness. The hope is that early treatment will impact long-term course and the subsequent development of disability, but there is, as yet, little convincing evidence that our current agents affect this outcome (Noseworthy, 2007). In addition, the currently available therapies are only partially effective, have side effects, are difficult to deliver, and are expensive. However, the widespread availability and clinical acceptance of these agents has led to a transformation in the design of modern MS clinical trials, one that is both ethically and practically based (McFarland and Reingold, 2005). Currently, more than ever, a dynamic pipeline of parenteral and oral agents is already in phase III testing so that several new agents may reach the market in the next few years. This new landscape of MS therapeutics presents novel challenges to future clinical trials, and this

chapter will review the assumptions and design considerations of pivotal and recent MS trials to provide a historical perspective on how we have arrived at the present moment in considering the future of MS research. It will conclude with an evaluation of the current state of ethics of placebo-controlled trials, as well as an overview of new approaches to the study of MS that take a more holistic approach than that of the traditional clinical trial.

CLINICAL TRIAL OUTCOMES MEASUREMENT: AN OVERVIEW

Multiple sclerosis clinical trials must be designed to capture the broad array of potential disease manifestations across individuals, but they must do so in a way that is reproducible and standardized. Outcome measures must be multidimensional in order to adequately encompass the myriad ways MS affects patients both in the short and long term. To this end, clinical trials focus on the two hallmark characteristics of MS: the occurrence of relapses and the accrual of disability. Choice of the outcome measure depends on the presumed mechanism of action of the investigated treatment and its anticipated clinical effect. It is important to choose the most appropriate primary outcome measure for each individual trial (D’Souza et al., 2008). In addition, a study must be of sufficient duration to allow the benefit of the agent to become evident and have a subject population large enough to power the study adequately. As long-term disability cannot be adequately assessed directly in a short-term clinical trial, all of our clinical measures from relapse-based assessments to measures of sustained disability in the short term can be considered surrogate markers of our ultimate long-term therapeutic goals.

Short-Term, Relapse-Based Outcomes

Clinical trials of disease-modifying agents for MS typically utilize relapse-based endpoints to demonstrate therapeutic effect. As short-term trials (usually between 1 and 3 years in duration) are often underpowered to demonstrate effect on long-term disability, endpoints such as the annualized relapse rate, time to first relapse, and percent of patients relapse free serve as surrogate

markers of disease suppression. These trials, and in particular their placebo arms, provide an instructive data source to characterize the short-term behavior of relapsing MS.

A relapse may be operationally defined as an objectively, clinically ascertainable, new or seriously worsening neurological deficit that persists for more than 24 hours, that develops at an interval of at least 30 days after a previous relapse, and is not related to infections. This 1-month window does not necessarily correspond with the actual biology of the disease, where often more than one area of active inflammation in the central nervous system (CNS) exists, each of which runs an independent time course (D'Souza et al., 2008). Short-lived aggravation of the symptoms related to elevations in body temperature (e.g., with fever, physical exercise, hot showers, warm weather) are referred to as Uthoff phenomena or "pseudorelapses." They result from an unmasking of subclinical lesions or worsening of chronic symptoms by transient elevations in temperature, and they do not constitute new inflammatory CNS activity. Patients cannot always distinguish a true exacerbation from a pseudorelapse, which is significant given that many trials rely on patients to report the occurrence of a relapse.

Problems comparing the efficacy of various treatments are amplified by trial-to-trial variation in relapse definition in terms of duration (24 or 48 hours) and whether a change on the neurological exam (as noted by a blinded evaluator) is present. Furthermore, trials are inconsistent about whether the objective change in the neurological examination must correspond with the clinical symptoms as described by the patient. The more narrow the definition, the lower the recorded relapse rate. Additionally, changes on the neurological examination (particularly changes in reflexes or subtle sensory changes) may be noted by the examiner in the absence of new symptoms to qualify as a relapse. Even with the most rigorous relapse criteria, borderline events will always be left to the investigators' judgment to interpret. If the expectations of the trial's sponsor are apparent to the investigators and their patients, the relapse rate in one group of patients might be underestimated in nonblinded studies (Sørensen, 2008).

Beyond the issues of ascertainment, relapse rate is a useful, but flawed, measure of MS disease activity.

Suppressing relapses benefits patients both immediately in terms of functional status, and potentially in the long term, as relapses are associated with significant residual deficits (Lublin et al., 2003). In addition, the relapse rate early in the disease is an important determinant of accumulation of disability later in the disease course. Although lower relapse rates in the early years of the disease may portend less eventual disability, whether a reduction in relapse rate imposed by treatment provides the same effect on long-term disability as would a relatively relapse-free experience as part of an individual's natural history remains unknown (Noseworthy et al., 2006).

Long-Term Observational Studies and Disability-Based Outcomes

The maximum clinical variability of MS is seen in the short term after disease onset; the illness becomes more predictable upon long-term observation. Much of the data on long-term outcomes in MS comes from several longitudinally followed cohorts and not from randomized clinical trials. As natural-history studies, these are largely comprised of patients who did not receive disease-modifying treatments; because of the availability of treatment, it is no longer feasible to follow such cohorts prospectively in the modern era. Three cohorts that provide large-scale, prospective analyses of untreated populations are those from Lyon, France (Confavreux et al., 2000), Sweden (Runmarker and Andersen, 1993), and London, Ontario (Weinshenker et al., 1991). Long-term prognosis is often described using time to development of disability landmarks in the Expanded Disability Status Score (EDSS). These include EDSS 4, the development of moderate disability with preserved gait; EDSS 6, the need for assistance with ambulation; and EDSS 7, the need for a wheelchair for mobility. While these classifications are limited by their focus on mobility, they provide a meaningful window into the time course of accrual of disability and reflect particular outcomes of great concern to patients diagnosed with MS.

The EDSS remains the criterion standard by which disability is measured in both the long and the short term. In short-term studies lasting from 1 to several years, "sustained disability" as

an outcome variable requires a measured change on the EDSS being reproduced at two points separated in time, usually by 6 months. If a patient “changes by 1 point on the EDSS,” however, this could mean a variety of things depending on where along the scale the patient is, and vary even further depending on what functional system brought about the change in EDSS step. In addition, persistent or sustained increases in physical impairment may occur because of a failure to completely recover from a relapse or from the slow accrual of disability that characterizes progressive MS. The use of “sustained disability” as a clinical endpoint fails to elucidate this distinction. As these processes are likely based on varying pathophysiology, lumping them together as a single clinical endpoint may not adequately capture the biologic response to a therapeutic agent.

The EDSS has a number of limitations that bear mention when considering its use in clinical trials. It is weighted toward ambulatory disability but is insensitive to other aspects of MS-related impairment, in particular cognitive dysfunction. Although a numerical scale, it is ordinal and categorical in nature and neither quantitative nor linear. The EDSS is thus not ideally suited to the deltas and changes in the “mean” that are frequently used as outcome variables. As an assessment tool, the EDSS has only moderate inter-rater reliability particularly at the lower range, is not entirely objective, and can have great fluctuation particularly at the lower levels. In its highest levels, the scale becomes a subjective description of a patient’s home care needs, and thus patients in the upper strata of EDSS are rarely included in clinical trials.

ISSUES IN CLINICAL TRIAL DESIGN AND INTERPRETATION

Reconsidering the Evolution of Pivotal Trials

The 1990s saw the publication of large clinical trials evaluating the three brands of interferon beta and glatiramer acetate as disease-modifying drugs in MS, and their subsequent regulatory approval and evolution into standard of care. Although the individual disease-modifying agents are discussed in previous chapters, some of the implications of

the pivotal trials on the design and interpretation of subsequent and recent studies will be reviewed here.

The randomized, placebo-controlled paradigm of the pivotal trials demonstrated not only efficacy of the agents studied but pronounced placebo effects on relapse rates when comparing prestudy with on-study exacerbation rates in the placebo group (D’Souza et al., 2008). The results of the interferon β -1b pivotal trial reported in 1993 (IFN β Multiple Sclerosis Study Group, 1993) are an archetype for this period in MS clinical research, and they demonstrated an annualized relapse rate of 0.78 in the interferon group versus 1.27 in the placebo group. The study was criticized for issues of relapse-ascertainment and confirmation, as relapses were self-reported and not universally confirmed by an examining neurologist. Three years later, in the weekly interferon β -1a study (Jacobs et al., 1996), the placebo group had an annualized relapse rate of 0.82, which was comparable to that of the rate of 0.84 in the placebo group of the pivotal trial of glatiramer acetate (Johnson et al., 1995). It was already clear that although there was a comparable degree of relative relapse rate reduction between the agents, the actual annualized relapse rates varied considerably between trials. All the values are likely to be skewed toward higher attack rates, as patients were selected for these studies on the basis of high prestudy clinical disease activity. The relapse rate may then “regress towards the mean” once these patients have been enrolled in the trial, in part explaining the extent of the apparent placebo effect. The efficacy of placebo may also reflect the impact of the comprehensive care provided to trial participants (D’Souza et al., 2008). The pivotal trials’ placebo group results underscore the importance of randomized, placebo-controlled trials and why positive head-to-head equivalency studies alone cannot be used to prove an agent’s efficacy, nor have they been considered sufficient evidence for regulatory approval.

The interferon β -1b pivotal trial did not demonstrate a statistically significant effect on disability, and as the trial enrolled ambulatory, relapsing-remitting patients it was likely underpowered to assess this outcome. In contrast, the pivotal trial of weekly intramuscular interferon β -1a (Jacobs et al., 1996) utilized disability as the primary

outcome measure and demonstrated a slower rate of accumulation of disability for the treated group, which was defined in advance as deterioration by ≥ 1 point on the EDSS for ≥ 6 months. In the subsequent pivotal trials of natalizumab, AFFIRM (Polman et al., 2006) and SENTINEL (Rudick et al., 2006), disability endpoints were also met, with sustained disability defined as persisting for ≥ 3 months.

A significant issue in the pivotal trials was that of the success of the blinding. Type and degree of blinding play an important role in the adjudication of patient-reported symptoms and their ascertainment by an evaluator. As the majority of treatments tested including glatiramer acetate, interferons, and mitoxantrone have easily recognized side effects or hallmarks such as injection site reactions, it is likely that the patients were able to correctly guess whether they were receiving placebo or active drug, which can confound the results given the subjective nature of relapses as assessed in these trials. Unscheduled symptom-initiated study visits and the methods of ascertainment of relapses at these visits should be equal across treatment arms to ensure that relapse assessment is not affected by systematic ascertainment bias. Modern double-blind studies have employed both a treating-neurologist and blinded evaluating neurologist to conduct the trial assessments. However, as oral and parenteral agents now in development have distinct modes of administration from the self-injected therapies with which they are being compared, this has necessitated using single-blind designs where maintenance of evaluator blinding is of paramount importance. Nonetheless, the lack of a double-blind design must be considered when interpreting these trials, which include the phase II and III trials of alemtuzumab (Coles et al., 2008).

Extension Trials and Long-Term Observational Studies

Given that clinical trials must establish efficacy in a short time frame, extension trials and open-label follow-up can provide a greater window into the benefit and safety profile of a therapeutic agent in the long term. In addition, there are ethical reasons to design an extension study, as it ensures continued access of an agent to the study population.

Typically, at the conclusion of a randomized controlled study, patients in the treatment group continue on the study drug, and patients initially randomized to the control group are offered the active drug and continue to be followed, albeit usually with a reduced frequency of assessments than during the randomized phase of the trial.

Although extension trials are frequently cited as evidence of the long-term efficacy of an agent, the quality of the data obtained in such an investigation is clearly inferior to that of the randomized phase, as the extension is open-label, uncontrolled, and the study population may exhibit self-selection by the patients who responded to the drug in the initial trial. Thus, extension trials may exhibit a significant selection bias toward “responders” and are of limited value in obtaining data that can be generalized overall regarding drug efficacy. In addition, extension trials are of limited impact with regard to outcome measures as they are generally no longer blinded, although to address this concern more recent trials have re-randomized the placebo group into a dose-blinded extension.

To provide even longer term assessments of the impact of a therapeutic agent, nonrandomized long-term observational studies have been undertaken lasting over a decade. These are in essence interminable extension trials, and their data suffer, at best, from similar methodological limitations. Large, long-term observational data sets for both the interferons and glatiramer have been analyzed (Ebers and Traboulsee, 2009; Rovaris et al., 2007) and are of greatest value in providing information about safety and long-term survival, while the “sustained efficacy” described in these data sets is confounded by enormous dropout rates, the unavoidable bias favoring responders, and the lack of an effective intention-to-treat analysis. These observational studies blur the distinction between research and clinical practice, and while they may attempt to quantify outcomes seen in “real world” standard of care, they lack the rigor and generalizability of well-performed clinical trials. That said, carefully designed long-term observational studies with specific hypotheses and preplanned analyses have the potential to provide valuable information that cannot be captured during the short-term randomized clinical trials of agents currently in testing.

The Problem of Cross-Trial Comparison

The reduction in relapse rate in the pivotal trial of weekly interferon β -1a (Jacobs et al., 1996) at 18% was more modest than that seen in the interferon β -1b pivotal trial. In the weekly interferon β -1a trial, the relapse rate declined from 1.2 to 0.61 in the treatment group versus 1.2 to 0.82 in the placebo group. These two trials underscore the problems of cross-trial comparison even for studies done in the same era, as the raw annualized relapse rate outcome would appear to favor weekly interferon β -1a as opposed to the 0.78 seen with interferon β -1b, while the relative reduction versus placebo favors the latter. Which drug is “better?” As discussed later in this chapter, this issue becomes all the more pronounced when one compares the results of the pivotal trials to those done in the “postmillennial” or “McDonald era” where an on-drug relapse rate of 0.78 or 0.61 would be considered an utter failure in the context of the current standard of efficacy.

Another example of the difficulties in comparing between two trials comes from the studies of the effect of subcutaneous interferon β -1b on disease progression in SPMS. The initial European study demonstrated that treatment with interferon β -1b was associated with a higher probability of stabilization of progression of disability compared to placebo. The therapeutic effect on this outcome measure, however, was not replicated in a North American study using a comparable but not identical study design (Panitch et al., 2004). Given the heterogeneity of MS in terms of clinical course and variable outcome, subtle variations in such factors as inclusion criteria, matching of study cohorts, selection of outcome variables, and statistical analyses are sufficient to render study results incomparable.

In the pivotal trial era, however, head-to-head studies were not performed. The Food and Drug Administration (FDA) does not accept “equivalency trials” as sufficient evidence for licensing approval, particularly as one could reasonably conclude from equivalence in a head-to-head study either that an agent is equally as effective or equally as ineffective as its comparator. This position favored the use of placebo-controlled trials, particularly as it statistically easier to demonstrate superiority of an investigational drug, with a

smaller sample size, when compared with an ineffective placebo than versus a partially-effective comparator.

The head-to-head trial era in MS began with the development of subcutaneous interferon β -1a given three times a week. As it is the same molecule as once-weekly interferon β -1ba, in order to expedite regulatory approval as Rebif, the FDA required positive head-to-head superiority studies before it could enter the market in the United States. Head-to-head comparisons of the relative efficacy of high- and low-dose interferon β -1a used as their relapse-based outcome the proportion of relapse-free and MRI activity-free patients. While both the INCOMIN (Durelli et al., 2002) and EVIDENCE studies (Panitch et al., 2002) demonstrated an advantage of high-dose interferon, by choosing as the primary outcome measure the “proportion of relapse-free patients” rather than the overall relapse rate, these trials evaluate the number of optimal responders and may not reflect a differential of general effectiveness across the population. They nonetheless were sufficient to justify the licensure of subcutaneous interferon β -1a in the United States, and they established a head-to-head trial paradigm that not all subsequent such studies were able to replicate.

The McDonald-Era Clinical Trials: Head-to-Head and Clinically Isolated Syndrome Trial Designs

Head-to-head studies

The successful implementation of active-comparator trials, including the head-to-head studies of high- versus low-dose interferon in the EVIDENCE and INCOMIN trials, led to the design and recruitment of several head-to-head trials of high-dose interferon β versus glatiramer acetate (Copaxone). Recruiting during the McDonald era, the BEYOND and REGARD trials (Mikol et al., 2008; see also Achiron and Fredrikson, 2009) failed to show a difference on relapse-based outcomes between either of the interferon β products and glatiramer acetate. These trials differ considerably in design, but the common factor involved in both trials’ failure to reach their relapse-based primary endpoints was the low event rate observed

in all groups. The REGARD study was a 2-year, randomized, open-label, head-to-head comparative study of subcutaneous interferon β -1a 44 μ g three times a week and glatiramer acetate, which showed no difference in the primary outcome measure, time to first relapse, or the proportion of patients who were relapse free. The sample size and power calculations, as previously described in the example above, are predicated on assumptions of an expected event rate. Although the annualized relapse rates in the REGARD study were almost identical (0.30 for interferon β -1a and 0.29 for glatiramer acetate), they were much lower than those reported in landmark trials (0.87 for interferon β -1a and 0.59 for glatiramer acetate) (Sørensen, 2008). The revised McDonald criteria allowed for the use of MRI to establish the diagnosis of MS; since many lesions detectable on MRI are clinically silent, these revised criteria allow a diagnosis of confirmed MS to be made earlier than would have been possible using clinical manifestations of relapses alone (Lublin, 2005). The inclusion of this new population in the postmillennial trials shifted the curve toward patients less likely to experience relapses in the short term.

A second consideration in the REGARD trial was the selection of time to first relapse as the primary outcome variable. Time to first relapse on treatment is a robust parameter, even in trials with high dropout rates. However, it does not make use of the second and subsequent relapses in the course of a trial, and thus it is particularly sensitive to differences in time course of therapeutic onset. It favors drugs with rapid onset, as compared to those with a more delayed but perhaps in the long run an equivalent or better-sustained effect (D'Souza et al., 2008). There are competing influences on trial design, where the goal is to assess a meaningful benefit in as little time as possible, but to be done in such a way as to demonstrate efficacy for a product intended for long-term use in a lifelong disease. It is of particular note that REGARD failed to demonstrate the purported more rapid time of onset of interferon β over glatiramer acetate.

Clearly the use of historical controls, particularly in the modern era of across-the-board decrease in observed relapse rates, imparts an unacceptably high degree of false-positive error. But what are the other implications of the relapse rates of the

McDonald era on clinical trial design and interpretation? One consideration is how to interpret the results of the AFFIRM trial (Polman et al., 2006) conducted at the outset of the McDonald era, which demonstrated the 68% reduction in relapse rate that natalizumab (Tysabri) established, approximately twice as great a reduction as the existing injectable therapies achieved in their pivotal trials. These same injectable agents, however, demonstrated approximately an 80% reduction in relapse rate in the BEYOND and REGARD trials of the past few years. Is AFFIRM the last trial of the pivotal era or the first of the McDonald era? There have been no head-to-head studies to elucidate this.

Either way, for most patients with MS, relapses are a relatively infrequent event. In studies planned for short duration, only a small percentage of patients will experience a relapse. It may be difficult to ascertain which of these patients respond optimally and which experienced either a more mild form of the disease, or just a period of disease quiescence unrelated to the therapeutic intervention (Walton, 2007). If most patients with RRMS experience either one or no relapses during a 2-year trial, the trial may be underpowered because of the low event rate, and it further calls into question the utility of the relapse outcome measure in trials of this duration. To maximize the expected event rate, some trials have employed clinical or radiographic entrance criteria (i.e., recent pre-enrollment relapses or enhancing lesions, respectively) during a specified time period. While this technique may increase the chances that an investigational drug will demonstrate a significant impact on a primary outcome variable, it further constrains the trial participants in a manner that further reduces the broad application of trial outcomes across the MS population and, of course, makes the trial more difficult to enroll.

In the pivotal trials of the 1990s, MRI outcomes were utilized to supplement the clinical data and provide further evidence supporting the drug's effectiveness. This was done to alleviate concerns about subjectively influenced efficacy findings such as effect on the relapse rate (Walton, 2007). Agents being tested in the treatment era have a higher standard of clinical efficacy to achieve and, to make matters more difficult, lower observed relapse rates on which to demonstrate their benefit.

Thus, surrogate markers like MRI are now utilized not to prove overall efficacy, which must be demonstrated clinically and convincingly, but rather to suggest a study subpopulation that may have a greater response. Magnetic resonance imaging and other surrogate markers can be used either as selection criteria or as a planned subgroup analysis to improve the risk–benefit ratio of an agent by identifying the population with the maximal response. The results of BEYOND and REGARD may have established clinical equipoise between interferon β and glatiramer acetate; however, these were both open-label, single-blind trials. We will have to wait for CombiRx, a large, ongoing, randomized, double-blind, double-dummy study that compares intramuscular interferon β -1a given weekly, glatiramer acetate given daily, and the combination of the two drugs given together to provide the highest-possible quality head-to-head evaluation of these agents (Sørensen, 2008). The CombiRx trial will also provide biomarker data to better identify patient characteristics that predict response to one of the therapeutic two modalities.

Clinically isolated syndrome trials

From a design perspective, clinically isolated syndrome (CIS) trials can be considered as RRMS trials recruiting after the first episode, where the outcome measures based on relapse occurrence are recast as time to “conversion to MS.” These trials have been implemented largely to satisfy regulatory requirements for an approved indication in CIS. Just as use of MRI in the McDonald criteria (McDonald et al., 2001; Polman et al., 2005) has allowed for the determination of dissemination in time sooner than would be found with clinical episodes alone, so, too, has the use of MRI in CIS trials provided radiographic outcome measures that dramatically raise the event rate of “conversion to MS.”

Building on the knowledge that MRI lesions at the time of a first demyelinating event are predictive of recurrence, large clinical trials of CIS patients have required clinically silent lesions as entrance criteria. Examining the data from the placebo arms of these trials lends support to the capacity of MRI lesions to predict conversion to MS. These trials include the Optic Neuritis Treatment Trial (ONTT), Controlled High Risk Subjects Avonex

Multiple Sclerosis Prevention Study (CHAMPS), the Early Treatment of Multiple Sclerosis Study (ETOMS), and more recently, the study of interferon β -1b in CIS, known as BENEFIT. The cumulative probability of developing a second attack in 2 years was 18% in the ONTT (Beck et al., 1995), 38% in CHAMPS (Jacobs et al., 2000), and 45% in ETOMS (Comi et al., 2001). The increased rates of conversion to MS in CHAMPS and ETOMS are attributable to the presence of clinically silent lesions on MRI as an enrollment criterion, underscoring the utility of initial MRI appearance to predict short-term outcomes. In BENEFIT (Kappos et al., 2006), after 2 years, 45% of placebo patients had experienced a second attack, but 85% fulfilled the McDonald criteria, demonstrating the increased sensitivity conferred by the radiographic diagnostic criteria and underscoring the hazards of evolving diagnostic criteria for cross-trial comparisons. Glatiramer acetate and intramuscular weekly interferon β -1a and interferon β -1b have demonstrated efficacy and garnered FDA approval for use at CIS, as would be anticipated from their established effect in relapsing remitting disease. A very low weekly dose of subcutaneous interferon β -1a also resulted in a statistically significant reduction in relapse rate, but this preparation is not currently approved by the FDA for the treatment of CIS.

Progressive Multiple Sclerosis and the Emerging Focus on Neuroprotection

In SPMS greater uncertainty exists regarding the short-term benefits of treatment; and the different results between the European and North American IFN trials in SPMS (Panitch et al., 2004) highlight the role that continued relapses and inflammatory activity play in the extent to which treatments have a “disease-modifying” effect on this form of the disease. This also highlights the continuum that exists between RRMS and SPMS, and how, even with restrictive inclusion criteria, trial populations differ in their range of pathophysiology such that one may reach its endpoint while the other fails to do so.

In addition, even those SPMS trials that do reach disability endpoints may not reflect the true risk–benefit profile of the agent. For example, the European trial of mitoxantrone in SPMS

(see Martinelli et al., 2005) demonstrated a decreased number of patients with EDSS progression, but at 188 patients it was too small to adequately predict the rare but serious potential consequences that have come to limit its use, including secondary leukemias. The issue of estimating such rare but potentially lethal side effects of the disease-modifying therapies was highlighted by the emergence of PML with natalizumab after the SENTINEL trial was completed. While this was initially estimated to be a “one in a thousand” risk potentially confined to its use in combination with interferon, the occurrence of PML in natalizumab monotherapy in post-marketing is a salient example of the difficulty of basing these estimates on the comparatively small population of a clinical trial.

Negative trials also provide valuable information for our understanding of the protean manifestations of the disease. The PROMiSe study, evaluating the effect of glatiramer acetate on accumulation of disability in PPMS, failed to reach its endpoint despite a sample size of over 900 patients and was halted after planned interim analysis. Akin to the methodological problem that caused REGARD and BEYOND to fail to reach their endpoints, the observed annual progression rate on the Expanded Disability Status Scale was only 36%, which was markedly lower than the 50% annual progression rate estimate used in the power calculations for sample size (Kieseier et al., 2007). Future clinical trial methodologies in progressive forms of MS are likely to focus on neuroprotection, which refers to targeting mechanisms of neuronal injury to prevent destruction, rather than the upstream efforts at immunomodulation that have shown minimal effect in progressive disease (Kapoor, 2006).

IMPLICATIONS FOR CURRENT AND FUTURE STUDY DESIGNS

Oral and Parenteral Agents

The goals of current MS clinical trials have evolved based on the successes that have preceded them. The pivotal trials established efficacy of the existing agents and assured their regulatory approval. These were followed by trial designs that confirmed

the utility of treatment at first event and subsequently by direct head-to-head and trials of comparative dosing strategies. Agents currently in clinical trials will be judged not only on the basis of reaching their study endpoints but also on what they add to the current armamentarium. To this end, there are several oral and parenteral drugs that are in phase III testing, and while it is as-yet unknown which will come to fruition, there are several paradigm changes in the current research era that bear mention.

Alemtuzumab is a monoclonal antibody given by annual cycles of IV infusion for which the results of a successful phase II trial were published in October 2008 (Coles et al., 2008). The alemtuzumab study was a bold design in two ways: instead of a placebo-control, the investigators used an active-comparator design, comparing two doses of alemtuzumab against subcutaneous interferon β -1a; and they chose sustained accumulation of disability as the primary endpoint. As discussed earlier, both of these design decisions decrease the likelihood of a positive trial, particularly when studied in a typical phase II sample size of 334 patients. Nonetheless, compared to interferon β -1a, relapses were reduced by 75% and sustained disability by 60% in patients receiving alemtuzumab. Importantly, the positive effects persisted at comparable magnitudes after 3 years (Kleinschnitz et al., 2008). These results have provided a considerably better understanding of how the novel treatment compares to current standard of care. While placebo-controlled trials have historically been the criterion standard, in the modern era comparator trials will be essential for clinicians and patients to make informed choices with regard to therapy (McFarland, 2009).

The encouraging results of the alemtuzumab trial were, however, tempered by six cases of idiopathic thrombocytopenic purpura (ITP), one with a fatal outcome leading to interruption of administration of study drug during the trial (Kleinschnitz et al., 2008). There were significant risks of autoimmune thyroid and renal diseases associated with this agent as well. Alemtuzumab can be viewed as a long-lasting form of selective immune ablation (Kleinschnitz et al., 2008), but the emergence of secondary autoimmune conditions may indicate that more powerful immunomodulatory agents aimed at acting more incisively on MS may entail

unwanted life-threatening perturbations of the immune system elsewhere, as was previously seen with the occurrence of PML with natalizumab (Lublin, 2005). The phase III alemtuzumab trials known as CARE-MS are currently ongoing, and they are slated to enroll approximately 2000 patients. They follow the same general paradigm as in phase II, comparing alemtuzumab against subcutaneous interferon β -1a; however, superior efficacy is all but a foregone conclusion on the basis of the phase II results. The phase III trials will thus largely serve to evaluate the incidence of adverse events such as ITP and autoimmune thyroid disease in a much larger study population. Unlike the pivotal phase III trials of the 1990s that were intended to establish an agent's clinical benefit, these phase III trials will enhance our ability to estimate the amount of risk in alemtuzumab's risk-benefit profile.

Rare but potentially lethal opportunistic infections have also been reported in the ongoing phase III studies of FTY720, or fingolimod, one of the promising oral agents in development. As new agents emerge that address the limited modes of administration and effectiveness of the current generation of therapeutics, this new potential for serious adverse effects mandates very close monitoring. With 15 years of experience with the current FDA-approved medications, physicians will face difficult decisions about whether and how to use new drugs that may be more efficacious but may also carry greater risks. New drugs in development must be designed with the understanding of the relatively modest benefit, but very substantial safety profile of the current era in MS therapeutics in mind.

Ethical and Practical Challenges to Placebo-Controlled Clinical Trials

While randomized, placebo-controlled clinical trials have been the mainstay and criterion standard for testing new drugs in MS, there are several challenges facing this design paradigm in the current era (McFarland, 2009). Because of the recent successes in developing new therapies for MS, placebo-controlled trials face three important hurdles: decreasing number of patients who qualify for studies; ethical concerns regarding the use of placebo when approved therapies are available;

and limited applicability of placebo-controlled results given the availability of several effective agents.

The ethics and future of placebo-controlled trials in MS have received considerable attention and have been the subject of a consensus opinion most recently revised in 2008. This resulted from a meeting of an international group of clinicians, ethicists, statisticians, regulators, and representatives from the pharmaceutical industry that convened under the auspices of the National MS Society's International Advisory Committee on Clinical Trials of New Agents in MS to reexamine the ethics of placebo-controlled clinical trials (Polman et al., 2008). The ethical conduct of placebo-controlled trials in MS rests upon the concept of clinical equipoise, which states that the lack of consensus among experts regarding preferred treatment justifies the need for clinical trials to further knowledge in the area. Even in the age of multiple approved drugs for MS, the condition still fits this definition, insofar as the medications are all incompletely effective, none halts relapses or progression, and all have problematic side effect profiles that affect tolerability and adherence (Polman et al., 2008). The number needed to treat (NNT) measure derived from published trials provides insight into the limitations of the current therapies: in patients with CIS, seven must be treated to prevent one patient from developing CDMS at 3 years. In RRMS, nine have to receive interferon to prevent a single relapse at 1 year, and eight must be treated for 2 years to prevent one patient from worsening by a single EDSS point during this interval (Noseworthy et al., 2006). Clearly much work remains to be done to discover and evaluate potentially better therapies. Despite these limitations, however, the use of placebo controls in clinical trials cannot be justified for those patients willing to initiate treatment with one of the approved medications (Miller, 2007).

The National MS Society Task Force (Polman et al., 2008) maintains that for patients with relapsing MS for which established effective therapies exist, placebo-controlled trials should only be offered with rigorous informed consent if the subjects refuse to use these treatments, have not responded to them, or if these treatments are not available to them for other reasons (e.g., economics) (Polman et al., 2008). Use of placebo-controlled

trials is ethical for patients whose disease falls outside the regulatory or regionally accepted criteria for treatment with available agents. This category has been taken to include PPMS, for which no treatment has been shown to be effective, and to some degree SPMS, for which there is limited consensus on the true risk–benefit ratio of the approved disease-modifying therapies (Miller, 2007).

Despite the requirements for “rigorous informed consent,” obtaining consent from a potential study subject may not be sufficient to ensure truly informed decision making. Contemporary informed consent documents are complex and voluminous. As an example, at the time of this writing the informed consent form for the current phase III trial of alemtuzumab is 13,000 words, or roughly the length of this chapter. Furthermore, these documents are often presented to patients with worsening disease, whose condition sometimes makes them feel greater pressure for potential trial participation (Polman et al., 2008).

Independent of the ethical issues, it is increasingly difficult to recruit and retain patients for placebo-controlled trials. Many recent trials have recruited subjects from Eastern European and Asian countries where treatment availability is variable; this raises further concerns regarding the provision of therapy to the participants at the conclusion of a trial. In addition, as discussed earlier with regard to the alemtuzumab trials, in this competitive treatment era, placebo-controlled trials fail to provide increasingly important information about comparative efficacy among available agents (McFarland and Reingold, 2005). Is it meaningful to know simply that a new agent is more effective than placebo when there are six agents already approved for the treatment of MS?

Alternatives to placebo-controlled designs

Particular focus has been placed on trials that either avoid the use of a placebo because of the ethical considerations or on designs that allow new therapies to be studied more rapidly or with fewer patients than would be needed in a conventional placebo-controlled trial (McFarland and Reingold, 2005). Such alternative trial designs include dose-finding strategies, where a placebo arm may be replaced by a lowest-dose arm not expected to confer maximal benefit but still conceptually an

improvement over placebo. Some have argued that it may be possible to conduct shorter, and therefore perhaps ethically and practically more acceptable trials. But the concept of a short-term study of an agent intended for a lifelong disease is controversial and even trials of 2 or 3 years' duration may be insufficient (McFarland and Reingold, 2005). Indeed, the major critique of the currently approved disease-modifying agents is that their long-term efficacy remains unproven.

Another alternative to monotherapy placebo-controlled trials are add-on trials, in which the investigational agent or placebo is added on to an existing therapy to evaluate whether the investigational drug will confer increased efficacy. This is best suited for use when the two agents have different and potentially synergistic mechanisms of action. Methodological issues arise owing to our often incomplete understanding of the drugs' molecular mechanisms of action, and consequent difficulty in predicting how the combination of agents will behave clinically. There is a potential for unexpected antagonism, as was recently seen in a small trial evaluating statins added on to interferon (Birnbbaum et al., 2008), as well as unexpected synergistic consequences, as with the incidence of PML when natalizumab was added on to interferon in the SENTINEL trial (Rudick et al., 2006). In addition, success of an agent in an add-on trial yields subsequent ethical and practical challenges when considering whether the new agent should be approved and prescribed as monotherapy if it was never tested in such a manner.

A further strategy is a deferred treatment/crossover design, where after an initial randomization to placebo or active drug, the placebo patients are switched to active treatment either at a prespecified interval or after they reach a clinical endpoint. This was the premise of the BENEFIT trial of interferon β -1b in CIS, whereby at conversion to CDMS the CIS patients on placebo were switched to interferon. This design addressed the ethical concerns inherent in keeping patients who now met McDonald criteria off of an approved therapy, and allowed for a comparison between early- and delayed-treatment in the CIS population. This trial was able to demonstrate both a significant risk reduction by interferon β on conversion from CIS to MS (Kappos et al., 2006), as well as a modest but significant decrease in risk of sustained

disability at 3 years (Kappos et al., 2007). If, however, one takes this study as incontrovertible evidence of the need for early treatment, it ethically precludes the further use of this placebo-controlled CIS design. This presents a paradigmatic example of the changing landscape of MS research: as existing treatments are shown to be superior to placebo in more numerous contexts, the study of new agents in the same manner is no longer ethically permissible.

As placebo-controlled trials can be performed with smaller populations than active-comparator or add-on trials, they potentially could expose fewer patients to a novel agent of unknown risk and unclear benefit than the larger trials required to test an investigational drug against an existing treatment. The deaths of patients in recent MS clinical trials of such agents as natalizumab (Goodin et al., 2008), alemtuzumab (Coles et al., 2008), and fingolimod (Leyboldt et al., 2009) are humbling reminders that every patient who agrees to participate in a trial of an experimental medication assumes not just foreseeable risks and those risks associated with placebo, but unforeseeable risks associated with novel therapeutics as well.

The conundrum in which the field of MS finds itself is that while randomized, controlled clinical trials are necessary to demonstrate efficacy on short-term, relapse-based outcomes, and are designed to conform to the standards of proof required by regulatory agencies for licensure, these 2–3 year trials are unlikely to ever answer the question as to whether long-term use of these agents dramatically alters the ultimate natural history of the disease. There are thus ethical and practical reasons to shorten the standard 2-year clinical trial paradigm while conceptual and theoretical reasons argue to study these agents long term.

Alternative outcomes for MS trials include composite outcome measures to reduce sample size and potentially shorten the duration of these studies. Composite outcomes comprise measures of multiple clinical domains and can thus increase the “event rate” as recorded in a trial. The EDSS and MSFC are themselves composite measures; however, by generating more comprehensive composite outcomes rather than a single primary outcome measure, the overall sensitivity of clinical measures can be increased. In addition, results derived from composite outcomes could potentially be more clinically meaningful than such single

outcome measures as “proportion of patients relapse free.” An example of an interesting possible future alternative to grading disability and mobility according to the artificial confines of the EDSS could be the use of a mobile device for long-term continuous measurement of physical activity. One such device, the ACTIBELT system, is currently being evaluated at the Sylvia Lawry Centre for Multiple Sclerosis Research. This wearable electronic device is integrated into the buckle of a clothing belt and provides quantitative information about patients’ physical activity under real-life conditions (D’Souza et al., 2008).

BEYOND THE CLINICAL TRIAL: CONSIDERATIONS FOR MULTIPLE SCLEROSIS HEALTH CARE AND ALTERNATIVE OUTCOMES

In addition to individual heterogeneity of symptoms, their severity, and their timing, the consequence of neurological impairment on a patient’s degree of functional disability varies considerably. A patient could be significantly neurologically impaired but not functionally disabled (e.g., a computer programmer who is wheelchair confined), while another can be vocationally disabled with minimal findings on neurological examination (e.g., a concert pianist with impaired right-hand proprioception). The former patient has an EDSS of 8, and the latter an EDSS of 2. Evaluating disability from a patient-centric perspective must take into account both objective impairment and its meaningful impact on the patient, including his or her coping strategies and quality of life. The focus in the field of MS has been for many years on the impact that investigational agents have on the commonly used rating scales; there has recently been a renewed emphasis on finding validated measures and performing well designed studies on our ability to improve patients’ lives.

Psychosocial, Quality of Life, and Vocational Impact Studies

Measuring the clinical impact of MS is a critical issue for judging experimental therapies in clinical trials. To this end, outcomes are being developed that take a more holistic view of patient well-being and response to a variety of modalities

of care beyond just those agents aimed at "modifying the disease." A formal investigative approach to the psychosocial consequences and emotional aspects of MS has become a major focus of research in recent years. The field of MS research has moved beyond the rating scales that are solely based on the neurological examination such as the EDSS, which are associated with weak psychometric properties. More patient-oriented measures like the Guy's Neurological Disability Scale (GNDS) and a wide range of disease-specific quality-of-life (QOL) instruments are characterized by high clinical relevance as they reflect the patient's perspective on the overall impact of medical intervention. Some of the examples of these metrics include the 36 Item Short Form Health Survey (SF 36), the Self-Rated Abilities for Health Practices Scale (SRAHP), the Health Promoting Lifestyle Profile II, and the MS Impact Scale (MSIS-29) (Amato and Portaccio, 2007).

The problems of access to care and services have also been increasingly addressed. Several studies provide evidence that MS patients benefit from rehabilitation and health promotion interventions. Exercise programs have been shown to improve strength, endurance, fatigue and functional abilities, mental and physical health, and health-related QOL. Access to these interventions is, however, oftentimes difficult for the physically disabled MS patients who need them the most. To overcome logistic barriers, a study by Bombardier and colleagues (2008) evaluated telephone counseling for Health Promotion in MS patients by using the HPLP II as a primary outcome measure and the SF 36 as a secondary outcome measure. The trial supports that telephone-based motivational interviewing is an effective method of health promotion in people with MS. Ennis and colleagues (2006) also found similar effectiveness of a health-promotion education program (the OPTIMISE program) among MS patients.

Evaluating the efficacy of rehabilitation interventions, however, is a challenging issue as patients in trials are not easily blinded and control groups are not readily established. Defining an outcome measure is of particular difficulty because of the individual variation of functional deficits and rehabilitation needs. As an example of the difficulty in proving the efficacy of rehabilitation, Storr and colleagues developed a trial methodology with a double-blinded design for evaluating the

short-term efficacy of multidisciplinary, inpatient rehabilitation among MS patients. Impairment was evaluated by the MSIS 29, and activity was assessed by the GNDS (Storr et al., 2006). Despite the use of the modern methodology, the study was unable to prove a significant advantage of rehabilitation for stable MS patients (Storr et al., 2006).

Vocational impact studies focus on the consequences of MS on patients during their employment years. Employment is not only an important factor in QOL of MS patients, but, by influencing aspects such as insurance status and means of access to health care, it can also be an important factor in their overall health (Wilson and Walker, 1993; Aronson, 1997). Early in the disease, patients with MS may be impaired in finding and maintaining a workplace not only due to obvious physical handicaps but also because of "hidden disabilities" such as subtle cognitive impairments, fatigue, coping problems, and even workplace discrimination (Johnson et al., 2004). Up to 50% of patients with MS lose their job, retire early, reduce their work hours, or change workplaces for less demanding but often also less-well-paid jobs (Kornblith et al., 1986; Jackson et al., 1991). This is especially a problem in persons suffering from a disease that in itself causes a major financial burden. A substantial part of the cost of MS relates to losses in earnings due to vocational problems (Whetten-Goldstein et al., 1998).

In a recent systematic review, Khan and colleagues (2009) reported that they found only two interventional studies focused on vocational rehabilitation for MS patients with enough internal validity to be regarded as meaningful enough to be included in their review. LaRocca and colleagues (1996) investigated job retention in employed MS patients, and Rumrill et al. (1998) focused on career entry in unemployed patients. Both of these studies still had major limitations such as small sample sizes, no blinding of participants, care providers, or assessors, and unclear intention to treat analyses. Nevertheless, they showed that vocational rehabilitation in MS was feasible and revealed that many patients tended to ignore or neglect possible future problems in finding and maintaining employment until they reached a crisis. Enabling persons with MS to achieve, maintain, or return to a status of employment and economic self-sufficiency is therefore one of the main aspects of rehabilitation, whether it is provided by general

MS rehabilitation or specialized MS vocational rehabilitation.

Multiple Sclerosis Databases and Resource Utilization Studies

The MS International Federation (MSIF) was established in 1967 as an international network targeted on a global linking of the activity of the national MS Society around the world. The Atlas of MS database by the MSIF provides a variety of data on the epidemiology of MS and the availability and accessibility of resources for the diagnosis, treatment, rehabilitation, and support for MS patients; it is aimed at improving life quality of people affected by MS. Another example of a useful database for MS is the New York State MS Consortium (NYSMSC), a regional alliance consisting of 15 MS centers throughout New York State. The NYSMSC developed a clinical long-term follow-up centralized MS patient registry, to offer a potential tool for interdisciplinary research. Similar to the MSIF Atlas, the NYSMSC database collection system is based on registry questionnaires and follow-up forms acquired at participating MS centers. In 1996, the North-American Consortium of MS Centers initiated the NARCOMS (North American Research Committee on Multiple Sclerosis) MS Patient Registry as a long-term project to promote and facilitate clinical and epidemiological research in MS. NARCOMS is a patient-driven, self-reported database with information collected semi-annually. Between 1996 and 2004, more than 28,000 participants were registered (Flachenecker and Stuke, 2008).

Resource utilization studies have examined the costs incurred by a diagnosis of MS, which is a topic of considerable importance to individual patients as well as the health-care delivery system in light of the annual expense of the disease-modifying agents. Kobelt and colleagues estimated the average total cost per American MS patient at \$47,000 per year in 2006, an increase of \$12,000 since 1995 (Kobelt et al., 2006). These numbers include direct costs such as medication and hospitalization, as well as indirect costs such as reduced working time and income or early retirement. Notably, the costs per person correlate very significantly with the number of relapses and stage of disability. Another study (O'Brien et al., 2003) determined the direct

medical cost of managing MS relapses in the United States, including the costs of hospitalization, ER, and physician office visits. The authors found that the majority of costs were attributable to hospital stays and concluded that the care of relapsing MS patients should be provided in the patients' homes as much as possible, while at the same time treatments that reduce the relapse frequency and thus the rate of hospital admissions may decrease patients' total financial burden.

Conclusions

The philosophy of early therapeutic intervention has begun to alter the natural course of MS, allowing patients to remain optimistic and play an active part in their disease management. Clinical trials for MS have been forever changed by the rapid development of disease-modifying agents, which raises both practical and ethical issues for the study of future agents targeted at the same outcomes. Clinical research has thus evolved to include both new agents and new outcome measures, including those that capture the impact of the disease in a patient-specific way not encapsulated by the standard neurologic exam-based metrics and arbitrary relapse definitions. As our current therapies do not result in an elimination of relapses or halt the accrual of disability, it is essential for a clinician to be aware of the widespread psychosocial impact of MS in terms of their patients' social involvement, partnership, family roles, and employment. The development of new effective agents for MS will require both a better biologic understanding of the disease's pathophysiology, likely with a focus on biomarkers, and the study of neuroprotection, as well as close attention to the risk/benefit profile of new agents based on an increasingly patient-centric approach. In the treatment era, there is an ongoing need for this work to broaden the perspective of neurologists and MS specialists on improving the care for patients with MS.

REFERENCES

- Achiron A, Fredrikson S. Lessons from randomised direct comparative trials. *J Neurol Sci.* 2009;277(suppl 1): S19-S24.
- Amato MP, Portaccio E. Clinical outcome measures in multiple sclerosis. *J Neurol Sci.* 2007;259(1-2):118-122.

- Aronson KJ. Quality of life among persons with multiple sclerosis and their caregivers. *Neurology*. 1997;48:74-80.
- Beck RW, Trobe JD. The Optic Neuritis Treatment Trial. Putting the results in perspective. The Optic Neuritis Study Group. *J Neuroophthalmol*. 1995;15(3):131-135.
- Birnbaum G, Cree B, Altafullah I, Zinser M, Reder AT. Combining beta interferon and atorvastatin may increase disease activity in multiple sclerosis. *Neurology*. 2008;71(18):1390-1395.
- Bombardier CH, Cunniffe M, Wadhvani R, et al. Counseling for Health Promotion in People With Multiple Sclerosis: A Randomized Controlled Trial. *Arch Phys Med Rehabil*. 2008; 89 (10):1849-1856.
- Coles AJ, Compston DA, Selmaj KW, et al. for the CAMMS223 Trial Investigators. Alemtuzumab vs. interferon beta-1a in early multiple sclerosis. *N Engl J Med*. 2008;359(17):1786-1801.
- Comi G, Filippi M, Barkhof F, et al. Early treatment of Multiple Sclerosis Study Group. Effect of early interferon treatment on conversion to definite multiple sclerosis: a randomised study. *Lancet*. 2001;357(9268):1576-1582.
- Confavreux C, Vukusic S, Moreau T, Adeleine P. Relapses and progression of disability in multiple sclerosis. *N Engl J Med*. 2000;343(20):1430-1438.
- D'Souza M, Kappos L, Czaplinski A. Reconsidering clinical outcomes in Multiple Sclerosis: relapses, impairment, disability and beyond. *J Neurol Sci*. 2008; 274(1-2):76-79.
- Durelli L, Verdun E, Barbero P, et al. Independent comparison of Interferon (INCOMIN) Trial Study Group. Every-other-day interferon beta-1b versus once-weekly interferon beta-1a for multiple sclerosis: results of a 2-year prospective randomised multicentre study (INCOMIN). *Lancet*. 2002;359.9316:1453-1460.
- Ebers GC, Reder AT, Traboulsee A, et al., Investigators of the 16-Year Long-Term Follow-Up Study. Long-term follow-up of the original interferon-beta1b trial in multiple sclerosis: design and lessons from a 16-year observational study. *Clin Ther*. 2009 Aug;31(8):1724-1736.
- Ennis M, Thain J, Boggild M, Baker GA, Young CA. A randomized controlled trial of a health promotion education programme for people with multiple sclerosis. *Clin Rehabil*. 2006;20(9):783-792.
- Flachenecker P, Stuke K. National MS registries. *J Neurol*. 2008;255(suppl 6):102-108.
- Goodin DS, Cohen BA, O'Connor P, Kappos L, Stevens JC, Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology. Assessment: the use of natalizumab (Tysabri) for the treatment of multiple sclerosis (an evidence-based review): report of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology. *Neurology*. 2008;71(10):766-773.
- IFNB Multiple Sclerosis Study Group. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. Clinical results of a multicenter, randomized, double-blind, placebo-controlled trial. *Neurology*. 1993;43(4):655-661.
- Jackson MF, Quaal C, Reeves MA. Effects of multiple sclerosis on occupational and career patterns. *Axone*. 1991;13:16-17, 20-22.
- Jacobs LD, Beck RW, Simon JH, et al. Intramuscular interferon beta-1a therapy initiated during a first demyelinating event in multiple sclerosis. CHAMPS Study Group. *N Engl J Med*. 2000;343(13):898-904.
- Jacobs LD, Cookfair DL, Rudick RA, et al. Intramuscular interferon beta-1a for disease progression in relapsing multiple sclerosis. The Multiple Sclerosis Collaborative Research Group (MSCRG). *Ann Neurol*. 1996;39(3):285-294.
- Johnson K, Amtmann D, Yorkston K, Klasner ER, Kuehn CM. Medical, psychological, social and programmatic barriers to employment for people with multiple sclerosis. *J Rehabil*. 2004;70:38-50.
- Johnson KP, Brooks BR, Cohen JA, et al. Copolymer 1 reduces relapse rate and improves disability in relapsing-remitting multiple sclerosis: results of a phase III multicenter, double-blind placebo-controlled trial. The Copolymer 1 Multiple Sclerosis Study Group. *Neurology*. 1995;45(7):1268-1276.
- Kapoor R. Neuroprotection in multiple sclerosis: therapeutic strategies and clinical trial design. *Curr Opin Neurol*. 2006;19(3):255-259.
- Kappos L, Freedman MS, Polman CH, et al. Effect of early versus delayed interferon beta-1b treatment on disability after a first clinical event suggestive of multiple sclerosis: a 3-year follow-up analysis of the BENEFIT study. *Lancet*. 2007;370(9585):389-397.
- Kappos L, Polman CH, Freedman MS, et al. Treatment with interferon beta-1b delays conversion to clinically definite and McDonald MS in patients with clinically isolated syndromes. *Neurology*. 2006;67(7):1242-1249.
- Khan F, Ng L, Turner-Stokes L. Effectiveness of vocational rehabilitation intervention on the return to work and employment of persons with multiple sclerosis. *Cochrane DB Sys Rev*. 2009;1:CD007256.
- Kieseier BC, Wiendl H, Hemmer B, Hartung HP. Treatment and treatment trials in multiple sclerosis. *Curr Opin Neurol*. 2007;20(3):286-293.
- Kleinschnitz C, Meuth SG, Wiendl H. The trials and errors in MS therapy. *Int MS J*. 2008;15(3):79-90.
- Kobelt G, Berg J, Lindgren P, Jönsson B. Costs and quality of life in multiple sclerosis in Europe: method of assessment and analysis. *Eur J Health Econ*. 2006; 7(suppl 2):S5-S13.
- Kornblith AB, LaRocca NG, Baum HM. Employment in individuals with multiple sclerosis. *Int J Rehabil Res*. 1986;9:155-165.
- LaRocca NG, Kalb RC, Gregg K. A program to facilitate retention of employment among persons with multiple sclerosis. *Work*. 1996;7:37-46.
- Leypoldt F, Münchau A, Moeller F, Bester M, Gerloff C, Heesen C. Hemorrhaging focal encephalitis under fingolimod (FTY720) treatment: a case report. *Neurology*. 2009;72(11):1022-1024.
- Lublin F. Multiple sclerosis trial designs for the 21st century: building on recent lessons. *J Neurol*. 2005;252: V46-V53.

- Lublin FD, Baier M, Cutter G. Effect of relapses on development of residual deficit in multiple sclerosis. *Neurology*. 2003;61(11):1528-1532.
- Martinelli Boneschi F, Rovaris M, Capra R, Comi G. Mitoxantrone for multiple sclerosis. *Cochrane DB Sys Rev*. 2005;4:CD002127.
- McDonald WI, Compston A, Edan G, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol*. 2001;50(1):121-127.
- McFarland HF. Alemtuzumab versus interferon beta-1a: implications for pathology and trial design. *Lancet Neurol*. 2009;8(1):26-28.
- McFarland HF, Reingold SC. The future of multiple sclerosis therapies: redesigning multiple sclerosis clinical trials in a new therapeutic era. *Mult Scler*. 2005;11(6):669-676.
- Mikol DD, Barkhof F, Chang P, et al. Comparison of subcutaneous interferon beta-1a with glatiramer acetate in patients with relapsing multiple sclerosis (the REBif vs Glatiramer Acetate in Relapsing MS Disease [REGARD] study): a multicentre, randomised, parallel, open-label trial. *Lancet Neurol*. 2008;7(10):903-914.
- Miller AE. Ethical considerations in multiple sclerosis clinical trials. In: Cohen JA, Rudick RA, eds. *Multiple Sclerosis Therapeutics*. 3rd ed. Boca Raton, FL: CRC Press; 2007: 343-356.
- Noseworthy JH. The challenge of long-term studies in multiple sclerosis: use of pooled data, historical controls, and observational studies to determine efficacy. In: Cohen JA, Rudick RA, eds. *Multiple Sclerosis Therapeutics*. 3rd ed. Boca Raton, FL: CRC Press; 2007: 319-330.
- Noseworthy JH, Miller D, Compston A. Disease modifying treatments in multiple sclerosis. In: Compston A, McDonald IR, Noseworthy J, Lassmann H, eds. *McAlpine's Multiple Sclerosis*. London: Churchill Livingstone Elsevier; 2006: 729-802.
- O'Brien JA, Ward AJ, Patrick AR, Caro J. Cost of managing an episode of relapse in multiple sclerosis in the United States. *BMC Health Serv Res*. 2003;3(1):17.
- Panitch H, Goodin DS, Francis G, et al. Evidence of interferon dose-response: European North American Comparative Efficacy; University of British Columbia MS/MRI Research Group. Randomized, comparative study of interferon beta-1a treatment regimens in MS - The EVIDENCE trial. *Neurology*. 2002;59(10):1496-1506.
- Panitch H, Miller A, Paty D, Weinshenker B, North American Study Group on Interferon beta-1b in Secondary Progressive MS. Interferon beta-1b in secondary progressive MS: results from a 3-year controlled study. *Neurology*. 2004;63(10):1788-1795.
- Polman CH, O'Connor PW, Havrdova E, et al. A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. *N Engl J Med*. 2006;354(9):899-910.
- Polman CH, Reingold SC, Barkhof F, et al. Ethics of placebo-controlled clinical trials in multiple sclerosis: a reassessment. *Neurology*. 2008;70(13 pt 2):1134-1140.
- Polman CH, Reingold SC, Edan G, et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". *Ann Neurol*. 2005;58(6):840-846.
- Rovaris M, Comi G, Rocca MA, et al. Long-term follow-up of patients treated with glatiramer acetate: a multicentre, multinational extension of the European/Canadian double-blind, placebo-controlled, MRI-monitored trial. *Mult Scler*. 2007;13(4):502-508.
- Rudick RA, Stuart WH, Calabresi PA, et al. Natalizumab plus interferon beta-1a for relapsing multiple sclerosis. *N Engl J Med*. 2006;354(9):911-923.
- Rumrill PD, Roessler RT, Cook BG. Improving career re-entry outcomes for people with multiple sclerosis: a comparison of two approaches. *J Voc Rehabil*. 1998;10(3):241-252.
- Runmarker B, Andersen O. Prognostic factors in a multiple sclerosis incidence cohort with twenty-five years of follow-up. *Brain*. 1993;116(pt 1):117-134.
- Sørensen PS. REGARD: what can we learn from randomised, open-label, head-to-head studies? *Lancet Neurol*. 2008;7(10):864-866.
- Storr LK, Sørensen PS, Ravnborg M. The efficacy of multidisciplinary rehabilitation in stable multiple sclerosis patients. *Mult Scler*. 2006;12(2):235-242.
- Walton MK. Selection and interpretation of end-points in multiple sclerosis clinical trials. In: Cohen JA, Rudick RA, eds. *Multiple Sclerosis Therapeutics*. 3rd ed. Boca Raton, FL: CRC Press; 2007: 295-308.
- Weinshenker BG, Rice GP, Noseworthy JH, Carriere W, Baskerville J, Ebers GC. The natural history of multiple sclerosis: a geographically based study. 3. Multivariate analysis of predictive factors and models of outcome. *Brain*. 1991;114(pt 2):1045-1056.
- Whetten-Goldstein K, Sloan FA, Goldstein LB, Kulas ED. A comprehensive assessment of the cost of multiple sclerosis in the United States. *Mult Scler*. 1998;4(5):419-425.
- Wilson SH, Walker GM. Unemployment and health: a review. *Pub Health*. 1993;107(3):153-162.